

Laborator 8

Săptămâna 13-17 aprilie 2020

1 Perceptron multistrat

Implementați algoritmul Backpropagation pentru o rețea feedforward multistrat de tip perceptron multistrat (MLP), pentru recunoaștere optică a caracterelor, folosind setul de date MNIST descris la adresa <http://yann.lecun.com/exdb/mnist/>, același ca la laboratorul 4.

MLP este prezentat în [1], capitolul 6.

Funcția de activare folosită pentru neuronii din stratul ascuns poate fi sigmoida logistică, tangenta hiperbolică sau ReLU. Neuronii din stratul de ieșire folosesc funcția de activare softmax, funcția de eroare este cross entropy.

Se va folosi regularizarea pentru ponderile legăturilor dintre noduri; ponderile asociate *bias*-urilor nu se regularizează. Se poate face învățare în stil gradient descent, minibatch sau batch (offline); recomandăm minibatch. Vor fi mai multe epoci de antrenare, până când valoarea funcției de eroare scade sub un anumit prag.

Inițializare valorilor ponderilor și a coeficienților de bias se face cu valori aleatoare mici în jurul lui 0; puteți opta de asemenea pentru strategia Glorot [3] sau He [4].

Scorul modelului antrenat se calculează ca acuratețe pe setul de testare din MNIST. Opțional, se poate reprezenta matricea de confuzie pentru setul de testare.

Se cere implementarea următoarelor două puncte:

1. Rețea neurală cu un singur strat ascuns, a se vedea figura 6.3 din curs; numărul de noduri de intrare este 784, numărul de neuroni din stratul ascuns se fixează apriori, e.g. 100, numărul de neuroni de ieșire este numărul de clase. Funcția de activare pentru neuronii din stratul de ieșire este softmax. Se raportează acuratețea de clasificare pe setul de antrenare și pe cel de testare.

2. Numărul de neuroni din stratul ascuns se determină prin **metoda hold-out**: setul de antrenare se partiționează în 50000 linii pentru antrenare, restul de 10000 pentru validare; rândurile din setul de instruire vor fi permutate în prealabil; pentru număr de neuroni din stratul se consideră o mulțime de valori candidat, de exemplu: $\{10, 50, 100, 200, \dots\}$. Se face antrenare folosind doar cele 50000 de date selectate ca mai sus, folosind mai multe epoci, după care se face testarea pe setul de validare. Numărul de neuroni din stratul ascuns pentru care eroarea pe setul de validare este cea mai mică se consideră a fi optimal și se folosește pentru a face antrenarea pe toate cele 60000 linii din setul de antrenare inițial (incluzând deci liniile din setul de validare); în final se raportează acuratețea pe seturile de (antrenare \cup validare) și respectiv de testare.

Opțional:

1. Se poate face **k-fold cross validation**, pentru a obține o evaluare mai veridică a performanțelor modelelor candidat. Un model candidat are o valoare anume pentru numărul de neuroni din stratul ascuns. Se poate lua $k = 10$.
2. Valoare adecvată a ratei de învățare se caută pe baza setului de validare.

Termen de predare: săptămâna 4–8 mai 2020.

Bibliografie

- [1] Lucian M. Sasu, *Inteligență artificială — note de curs*.
- [2] Michael Nielsen, *Neural Networks and Deep Learning*, <http://neuralnetworksanddeeplearning.com/>
- [3] Xavier Glorot, Yoshua Bengio, *Understanding the difficulty of training deep feedforward neural networks*, Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS'10). Society for Artificial Intelligence and Statistics, 2010
- [4] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, *Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification*, Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), 2015